

# Application of Information Theory, Lecture 12

## Accessible Entropy and Statistically Hiding Commitments

### Handout Mode

Iftach Haitner

Tel Aviv University.

January 05, 2016

# Section 1

## **Commitment Schemes**

# Motivation

- ▶ Digital analogue of a safe
- ▶ Numerous applications (e.g., zero-knowledge, coin-flipping, secure computations )

## Definition

$\mu$  is **negligible**, denoted  $\mu(n) = \text{neg}(n)$ , if  $\forall p \in \text{poly} \exists n' \in \mathbb{N}$  s.t.  $\mu(n) < \frac{1}{p(n)}$  for all  $n > n'$ .

### Definition 1 (Commitment scheme)

An efficient two-stage protocol  $(S, R)$ .

- ▶ Commit stage: The sender  $S$  has private input bit  $b \in \{0, 1\}$  and a common input is  $1^n$ . Let  $\text{trans}$  be the transcript of this stage.
- ▶ Reveal stage:  $S$  sends the pair  $(r, b)$  to  $R$ , and  $R$  **accepts** if  $\text{trans}$  is consistent with  $S(\sigma, r)$ .

**Hiding:** Let  $V_n^{R^*}(b)$  be  $R^*$ 's view in (the commit stage of)  $(S(b), R^*)(1^n)$ .

Then for **any**  $R^*$ :  $\Delta^{R^*}(V_n^{R^*}(0), V_n^{R^*}(1)) = \text{neg}(n)$ .

**Binding:** The following happens with negligible probability for **any**  $S^*$ :

$S^*(1^n)$  interacts with  $R(1^n)$  in the commit stage resulting in transcript  $\text{trans}$ . Then  $S^*$  outputs two strings  $r_0$  and  $r_1$  such that  $R(\text{trans}, r_0, 0) = R(\text{trans}, r_1, 1) = \text{Accept}$ .

**Alternative Binding definition:** Assume that following the interaction  $S^*$  outputs a pair  $(r, b)$  with  $R(\text{trans}, r, b) = \text{Accept}$ . Let  $V^{S^*}$  be  $S^*$ 's view in (the commit stage of)  $(S^*, R^*)(1^n)$ . Then  $H(b|V^{S^*}) = \text{neg}(n)$ .

## Definition cont.

- ▶ Naturally extends to strings
- ▶ Hiding: Perfect, statistical, computational.
- ▶ Binding: Perfect, statistical, computational.
- ▶ Impossible to have simultaneously both properties to be statistical.
- ▶ OWF is necessary assumption
- ▶ OWFs imply both statistically binding and computationally hiding commitments, and (more difficult) computationally binding and statistically hiding commitments.
- ▶ We focus on computationally binding, and statistically hiding commitments (SHC)

## Section 2

# Inaccessible Entropy

# Motivation

## Definition 2 (collision resistant hash family (CRH))

Function family  $\mathcal{H} = \{\mathcal{H}_n: \{0, 1\}^n \mapsto \{0, 1\}^{n/2}\}$  is **collision resistant**, if  $\forall$  PPT  $A$

$$\Pr_{\substack{h \leftarrow \mathcal{H}_n \\ (x, x') \leftarrow A(1^n, h)}} [x \neq x' \in \{0, 1\}^* \wedge h(x) = h(x')] = \text{neg}(n)$$

- ▶ Implies SHC. (?) Believed **not** to be implied by OWFs.
- ▶ Assume for simplicity that  $h \in \mathcal{H}_n$  is  $2^{n/2}$  to 1 and that a PPT cannot find a collision in any  $h \in \mathcal{H}_n$
- ▶ Given  $h(U_n)$ , the (min) entropy of  $U_n$  is  $n/2$ .
- ▶ Consider PPT  $A$  that on input  $h$  first outputs  $h, y$ , and then outputs  $x \in h^{-1}(y)$  (possibly using additional random coins)
- ▶ What is the entropy of  $x$  given  $(h, y)$  and the coins  $A$ 's used to sample  $y$ ? (essentially) 0!
- ▶ The generator  $G(h, x) = (h, h(x), x)$  has **inaccessible entropy**  $n/2$
- ▶ Does inaccessible entropy generator implies SHC?
- ▶ Does OWF implies inaccessible entropy generator?

# Real entropy

- ▶ Sample entropy: for rv  $X$  let  $H_X(x) = -\log \Pr_X [x]$ .
- ▶  $H(X) = \mathbb{E}_{x \leftarrow X} [H_X(x)]$
- ▶ For rvs  $X$  and  $Y$ , let  $H_{X|Y}(x|y) = H_{X|Y=y}(x)$ .
- ▶  $X_{1,\dots,i}$  stand for  $X_1, \dots, X_i$
- ▶ Let  $G: \{0, 1\}^n \mapsto (\{0, 1\}^\ell(n))^{m(n)}$  be an  $m$ -block generator
- ▶ Let  $(G_1, \dots, G_m) = G(U_n)$
- ▶ For  $\mathbf{g} = (g_1, \dots, g_m) \in \text{Supp}(G_1, \dots, G_m)$ , let

$$\text{RealH}_G(\mathbf{g}) = \sum_{i \in [m]} H_{G_i|G_1, \dots, G_{i-1}}(g_i|g_1, \dots, g_{i-1})$$

- ▶ The real Shannon entropy of  $G$ , wrt security parameter  $n$ , is

$$\mathbb{E}_{\mathbf{g} \leftarrow G(U_n)} [\text{RealH}_{G,n}(\mathbf{g})]$$

- ▶  $\mathbb{E}_{\mathbf{g} \leftarrow G(U_n)} [\text{RealH}_{G,n}(\mathbf{g})] = \sum_{i \in [m]} H(G_i|G_1, \dots, G_{i-1}) = H(G(U_n))$



## Accessible entropy

- ▶ Let  $G$  be an  $m$  block generator.
- ▶ Let  $\tilde{G}$  be an  $m$ -block generator, that uses coins  $r_i$  before outputting its  $i$ 'th block  $(w_i, g_i)$ .
- ▶  $t = (r_1, w_1, g_1, \dots, r_m, w_m, g_m)$  is **valid** with respect to  $G$ , and  $n$ , if  $w_i \in \{0, 1\}^n$  and  $(g_1, \dots, g_i) = G(w_i)_{1, \dots, i}$  for every  $i \in [m]$ .
- ▶ We assume for simplicity that  $t$  is **always** valid, and omit  $w$ 's.
- ▶  $\tilde{T} = (\tilde{R}_1, \tilde{G}_1, \dots, \tilde{R}_m, \tilde{G}_m)$ —the rv's induced by random execution of  $\tilde{G}(1^n)$

$$\begin{aligned}\text{AccH}_{\tilde{G}, n}(t) &= \sum_{i \in [m]} H_{\tilde{G}_i | \tilde{R}_1, \tilde{G}_1, \dots, \tilde{R}_{i-1}, \tilde{G}_{i-1}}(g_i | r_1, g_1, \dots, r_{i-1}, g_{i-1}) \\ &= \sum_{i \in [m]} H_{\tilde{G}_i | \tilde{R}_1, \dots, \tilde{R}_{i-1}}(g_i | r_1, \dots, r_{i-1})\end{aligned}$$

- ▶ The **accessible entropy** of  $\tilde{G}$  (wrt  $G$ ), and  $n$ , is at most  $k$ , if  $\Pr_{t \leftarrow \tilde{T}} [\text{AccH}_{\tilde{G}, n}(t) > k] \leq \text{neg}(n)$ . Why not  $\mathbb{E}_{t \leftarrow \tilde{T}} [\text{AccH}_{\tilde{G}, n}(t)]$ ?
- ▶ **inaccessible entropy**
- ▶ We will omit  $n$  when clear from the context

## Example

- ▶ Let  $\mathcal{H} = \{\mathcal{H}_n: \{0, 1\}^n \mapsto \{0, 1\}^{n/2}\}$  be  $2^{n/2}$ -to-1 collision resistant, and assume for simplicity that a PPT cannot find a collision for any  $h \in \mathcal{H}_n$ .
- ▶ Let  $G$  be the 3-block generator  $G(h, x) = (h, h(x), x)$
- ▶ Real entropy of  $G$  is  $\log |\mathcal{H}_n| + n$
- ▶ Accessible entropy of  $G$  is  $\log |\mathcal{H}_n| + \frac{n}{2}$

## Section 3

# Manipulating Inaccessible Entropy

## Entropy equalization

Let  $G$  be  $m$ -bit generator.

For  $\ell \in \text{poly}$  let  $G^{\otimes \ell}$  be the following  $(\ell - 1) \cdot m$ -bit generator

$$G^{\otimes \ell}(x_1, \dots, x_\ell, i) = G(x_1)_i, \dots, G(x_1)_m, \dots, G(x_\ell)_1, \dots, G(x_\ell)_{i-1}$$

- ▶ Assume the accessible entropy of  $G$  is (at most)  $k_A$ , then  $k_A^{\otimes \ell}$ , the accessible entropy of  $G^{\otimes \ell}$ , is at most  $k(\ell - 2) + m$ .
- ▶ Assume the real entropy of  $G$  is  $k_R$ , then

1. For any  $i \in [(\ell - 1) \cdot m]$  and  $(g_1, \dots, g_{i-1}) \in \text{Supp}(G_1^{\otimes \ell}, \dots, G_{i-1}^{\otimes \ell})$ :

$$H(G_i^{\otimes \ell} | G_1^{\otimes \ell}, \dots, G_{i-1}^{\otimes \ell}) \geq k_R / \ell$$

2.  $k_R^{\otimes \ell}$ , the real entropy of  $G^{\otimes \ell}$ , is at least  $(\ell - 1)K_R$

- ▶ Assume  $k_R \geq k_A + 1$ , then for  $\ell = m + 2$ , it holds that  $k_R^{\otimes \ell} \geq k_A^{\otimes \ell} + 1$

## Parallel repetition

Let  $G$  be an  $m$ -block generator and for  $\ell \in \text{poly}$ , let  $G^\ell$  be the  $\ell$ -fold parallel repetition of  $G$ .

- ▶ Assume accessible entropy of  $G$  is (at most)  $k_A$ , then the accessible entropy of  $G$  is at most  $k_A^\ell = \ell k_A$ .
- ▶ Assume  $H(G_i | G_1, \dots, G_{i-1}) = k_R$  for any  $i \in [m]$ , then for any  $i \in [m]$  and  $(g_1^\ell, \dots, g_{i-1}^\ell) \in \text{Supp}(G_1^\ell, \dots, G_{i-1}^\ell)$  it holds that

$$k_{\min}^\ell = H_\infty(G_i^\ell | G_1^\ell, \dots, G_{i-1}^\ell) \approx \ell k_R$$

- ▶ If  $k_A \leq k_R - 1$ , then  $\forall n \in \text{poly} \exists \ell \in \text{poly}$  such that  $\ell k_{\min}^\ell > k_A^\ell + n$

## Section 4

# Inaccessible Entropy from OWF

# The generator

## Definition 3

Given a function  $f: \{0, 1\}^n \mapsto \{0, 1\}^n$ , let  $G$  be the  $(n + 1)$ -block generator

$$G(x) = f(x)_1, \dots, f(x)_n, x$$

## Lemma 4

*Assume that  $f$  is a OWF then  $G$  has accessible entropy at most  $n - \log n$ .*

- ▶ Recall  $f$  is OWF if  $\Pr_{x \leftarrow \{0,1\}^n} [\text{Inv}(f(x)) \in f^{-1}(f(x))] = \text{neg}(n)$  for any PPT  $\text{Inv}$ .
- ▶ The real entropy of  $G$  is  $n$
- ▶ Hence, inaccessible entropy gap is  $\log n$
- ▶ Proof idea

## Proving Lemma 4

Let  $\tilde{G}$  be a PPT, and assume  $\Pr \left[ \text{AccH}_{G, \tilde{G}}(\tilde{T}) \geq n - \log n \right] \geq \varepsilon = \frac{1}{\text{poly}(n)}$ .

(recall  $\tilde{T} = (\tilde{R}_1, \tilde{G}_1, \dots, \tilde{R}_m, \tilde{G}_m)$  is the coins and output blocks of  $\tilde{G}$ )

### Algorithm 5 ( $\text{Inv}(z)$ )

1. For  $i = 1$  to  $n$ , do the following for  $n^2/\varepsilon$  times:
    - 1.1 Sample  $r_i$  uniformly at random and let  $g_i$  be the  $i$ 'th output block of  $\tilde{G}(r_1, \dots, r_i)$ .
    - 1.2 If  $g_i = z_i$ , move to next value of  $i$ .
    - 1.3 Abort, if the maximal number of attempts is reached.
  2. Finish the execution of  $\tilde{G}(r_1, \dots, r_{n+1})$ , and output its  $(n+1)$  output block.
- 
- ▶ We start by assuming that  $\text{Inv}$  is **unbounded** (replace  $n^2/\varepsilon$  with  $\infty$ )
  - ▶  $\hat{T} = (\hat{R}_1, \hat{G}_1, \dots, \hat{R}_{n+1}, \hat{G}_{n+1})$  is the (final) values of  $(r_1, g_1, \dots, r_{n+1}, g_{n+1})$  in a random execution of  $\text{Inv}(f(U_n))$ .



# $\tilde{T}$ vs. $\hat{T}$

- ▶ Fix  $\mathbf{t} = (r_1, g_1, \dots, r_{n+1}, g_{n+1}) \in \text{Supp}(\tilde{T})$
- ▶ Let  $P(\mathbf{t}) = \prod_{i=1}^{n+1} \Pr[\tilde{R}_i = r_i \mid (\tilde{R}_{1,\dots,i-1}, \tilde{G}_i) = (r_{1,\dots,i-1}, g_i)]$

$$\begin{aligned} \Pr_{\tilde{T}}[t] &= \Pr[\tilde{G}_1 = g_1] \cdot \Pr[\tilde{R}_1 = r_1 \mid \tilde{G}_1 = g_1] \cdot \Pr[\tilde{G}_2 = g_2 \mid \tilde{R}_1 = r_1] \\ &\quad \cdot \Pr[\tilde{R}_2 = r_2 \mid \tilde{G}_2 = g_2] \cdots \\ &= P(\mathbf{t}) \cdot \Pr[\tilde{G}_1 = g_1] \cdot \Pr[\tilde{G}_2 = g_2 \mid \tilde{R}_1 = r_1] \cdots \\ &= P(\mathbf{t}) \cdot 2^{-\sum_{i=1}^m H_{\tilde{G}_i \mid \tilde{R}_{1,\dots,i-1}}(g_i \mid r_{1,\dots,i-1})} \\ &= P(\mathbf{t}) \cdot 2^{-\text{AccH}_{\tilde{G}}(\mathbf{t})} \end{aligned}$$

- ▶  $\Pr_{\hat{T}}[\mathbf{t}] = \Pr[f(U_n) = g_{1,\dots,n}] \cdot \Pr[\tilde{G}_{n+1} = g_{n+1} \mid \tilde{R}_{1,\dots,n} = r_{1,\dots,n}] \cdot P(\mathbf{t})$
- ▶  $\Pr_{\hat{T}}[\mathbf{t}] = \frac{\Pr[f(U_n)=g_{1,\dots,n}] \cdot \Pr[\tilde{G}_{n+1}=g_{n+1} \mid \tilde{R}_{1,\dots,n}=r_{1,\dots,n}]}{2^{-\text{AccH}_{\tilde{G}, \tilde{G}}(\mathbf{t})}} \cdot \Pr_{\tilde{T}}[\mathbf{t}]$

## $\tilde{T}$ vs. $\hat{T}$ cont.

- ▶  $\mathbf{t} = (r_1, g_1, \dots, r_{n+1}, g_{n+1}) \in \text{Supp}(\tilde{T})$
- ▶  $\Pr_{\hat{T}}[\mathbf{t}] = \frac{\Pr[f(U_n)=g_1, \dots, n] \cdot \Pr[\tilde{G}_{n+1}=g_{n+1} | \tilde{R}_{1, \dots, n}=r_1, \dots, n]}{2^{-\text{AccH}_{G, \tilde{G}}(\mathbf{t})}} \cdot \Pr_{\tilde{T}}[\mathbf{t}]$
- ▶ Note that  $\Pr[f(U_n) = g_1, \dots, n] \cdot \frac{1}{|f^{-1}(g_1, \dots, n)|} = 2^{-n}$
- ▶ Hence, for  $\mathbf{t}$  with  $\text{AccH}_{G, \tilde{G}}(\mathbf{t}) \geq n - \log n$  and  $\Pr[\tilde{G}_{n+1} = g_{n+1} | \tilde{R}_{1, \dots, n} = r_1, \dots, n] \geq \frac{\alpha}{|f^{-1}(g_1, \dots, n)|}$ :

$$\Pr_{\tilde{T}}[\mathbf{t}] \geq \frac{\alpha}{n} \cdot \Pr_{\hat{T}}[\mathbf{t}] \quad (1)$$

## Inv's success probability

Let  $\mathcal{S} \subseteq \text{Supp}(\tilde{T})$  denote the set of transcripts  $\mathbf{t} = (r_1, g_1, \dots, r_{n+1}, g_{n+1})$  with

1.  $\text{AccH}_{\tilde{G}}(\mathbf{t}) \geq n - \log n$ ,
2.  $H_{\tilde{G}_i | \tilde{R}_1, \dots, i-1}(g_i | r_1, \dots, i-1) \leq \log(\frac{4n}{\varepsilon})$  for all  $i \in [n]$ ,
3.  $H_{\tilde{G}_{n+1} | \tilde{R}_1, \dots, n}(g_{n+1} | r_1, \dots, n) \leq \log(\frac{4}{\varepsilon} \cdot |f^{-1}(g_1, \dots, n)|)$ .

- ▶  $\Pr_{\tilde{T}} \left[ \exists i \in [n]: H_{\tilde{G}_i | \tilde{R}_1, \dots, i-1}(g_i | r_1, \dots, i-1) > \log(\frac{4n}{\varepsilon}) \right] \leq n \cdot \frac{\varepsilon}{4n} = \varepsilon/4$
- ▶  $\Pr_{\tilde{T}} \left[ H_{\tilde{G}_{n+1} | \tilde{R}_1, \dots, n}(g_{n+1} | r_1, \dots, n) > \log(\frac{4}{\varepsilon} \cdot |f^{-1}(g_1, \dots, n)|) \right] \leq \varepsilon/4$
- ▶  $\Pr_{\tilde{T}}[\mathcal{S}] \geq \Pr[\text{AccH}_{G, \tilde{G}}(T) \geq n - \log n] - 2 \cdot \frac{\varepsilon}{4} \geq \frac{\varepsilon}{2}$
- ▶ By Eq. (1):  $\Pr_{\hat{T}}[\mathcal{S}] \geq \frac{\varepsilon/4}{n} \cdot \Pr_{\hat{T}}[\mathcal{S}] \geq \frac{\varepsilon^2}{8n} \dots$

Back the **bounded** version of **Inv**.

- ▶ For  $z \in \{0, 1\}^n$  for which  $\exists (r_1, z_1, \dots, r_n, z_n, \dots) \in \mathcal{S}$ :  
 $\Pr[\text{Inv}(z) \text{ aborts}] \leq n \cdot (1 - \frac{\varepsilon}{4n})^{n^2/\varepsilon} = O(n \cdot 2^{-n}) \leq \frac{1}{2}$
- ▶ Hence,  $\Pr_{\hat{T}}[\mathcal{S}] \geq \frac{\varepsilon^2}{16n} \implies \Pr_{x \leftarrow \{0, 1\}^n}[\text{Inv}(f(x)) \in f^{-1}(f(x))] \geq \frac{\varepsilon^2}{16n}$

## Section 5

# **Statistically Hiding Commitment from Inaccessible Entropy Generator**

## High-level description

- ▶ Entropy equalization + gap amplification to get generator that has the **same** min-entropy in each block and whose accessible entropy is  $n$ -bit smaller than the sum of the min entropies.
- ▶ Use "hashing protocol" to get a "generator" with **zero** accessible entropy block
- ▶ Use a random block to mask the committed bit, to get a **weakly binding** SHC
- ▶ Amplify the above into full-fledged SHC